Method for investigating cytosine methylation in DNA sequences by means of hemi-methylation sensitive restriction enzymes

The present invention concerns a method for investigating cytosine methylation in DNA sequences.

Background of the invention

5-Methylcytosine is the most frequent covalently modified base in the DNA of eukaryotic cells. It plays an important biological role, among other things, in the regulation of transcription, in genetic imprinting and in tumorigenesis (for review: Millar et al.: Five not four: History and significance of the fifth base. In: The Epigenome, S. Beck and A. Olek, eds.: The Epigenome. Wiley-VCH Publishers, Weinheim, 2003, pp. 3-20). The identification of 5-methylcytosine as a component of genetic information is thus of considerable interest. The detection of methylation is difficult, since cytosine and 5-methylcytosine have the same base pairing behavior. Many of the conventional detection methods based on hybridization thus cannot distinguish between cytosine and methylcytosine. In addition, methylation information is completely lost during a PCR amplification.

The usual methods for methylation analysis operate according to two essentially different principles. In the first, a selective chemical conversion of unmethylated cytosines to uracil (bisulfite treatment) is employed, and in the second, methylation-specific restriction enzymes are used. DNA that has been pretreated enzymatically or chemically is then amplified and can be analyzed in different ways (for review: WO 02/072880, pp. 1 ff).

The conventional methods suffer from several disadvantages. For example, bisulfite treatment takes time and is labor-intensive. There is also the possibility that the DNA is only incompletely converted and also that it is partially degraded. Quantification is difficult both for the chemically as well as the enzymatically pretreated DNA. An amplification is necessary for this

purpose, usually by means of a PCR. This additional working step is associated with several problems, e.g. the risk of a preferred amplification of specific sequences (so-called "bias").

The method according to the invention, in contrast, does not require an amplification and thus permits a more rapid and simpler analysis than the conventional methodology. A quantification is also enabled. In the method according to the invention, the DNA to be investigated is hybridized to oligonucleotides of a defined methylation state. In this way, depending on the methylation status of the DNA to be investigated and on the oligonucleotides employed, hybrids are formed, which possess either the same or a different methylation state on each of the two DNA strands. Subsequently, the hybrids are reacted with restriction enzymes, wherein the restriction is dependent on the methylation state of the hybrids. Then the methylation status of the DNA can be concluded from the occurrence or non-occurrence of restriction which may be detected by means of various possible detection techniques known in the art.

Hybrid formation and a subsequent different restriction of the differently methylated hybrids is also utilized in the case of so-called Genomic Mismatch Scanning (GMS). This concerns a method for the detection of polymorphisms. Here, the DNA strands of two different individuals are hybridized to one another, wherein the DNA of one individual has been synthetically methylated beforehand with the use of enzymes. The homohybrids that are formed are then digested by enzymes, while the heterohybrids are further analyzed (see, e.g.: Nelson et al.: Genomic mismatch scanning: a new approach to genetic linkage mapping. Nat Genet. 1993 May; 4 (1) :11- 8). The synthetic methylation and the subsequent restriction thus are used here for the isolation of heterohybrids. The utilization of the different restriction of differently methylated DNA hybrids for the quantitative analysis of natural DNA methylation patterns has not been described previously.

Description

The use of the term "sensitive surface" shall be taken to mean a surface whose physical or chemical properties can be modified in a detectable manner wherein an oligonucleotide bound or otherwise present on said surface is modified, preferably by means of a restriction, digestion or degradation of said oligonucleotide.

The use of the term "hybridisation" shall be taken to mean hybridization under stringent or moderately stringent conditions. The stringency of hybridisation is determined by a number of factors during hybridisation and during any subsequent washing procedures, including temperature, ionic strength, length of time and concentration of formamide. These factors are outlined in, for example, Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2d ed., 1989).

The term "hemi-methylated" or "hemi-methylation" refers to a double stranded nucleic acid comprising at least one CpG position wherein the cytosine thereof is methylated on one strand only.

The term "homogenously methylated" or "homogenous methylation" refers to a double stranded nucleic acid comprising at least one CpG position wherein the cytosine thereof is methylated on both strands.

The term "unmethylated" when used in reference to a double stranded nucleic acid, single stranded nucleic acid, double stranded oligonucleotide or single stranded oligonucleotide comprising at least one CpG position shall be taken to mean that no cytosines of said CpG dinucleotides are methylated.

The term "methylated" when used in reference to a double stranded nucleic acid, single stranded nucleic acid, double stranded oligonucleotide or single stranded oligonucleotide comprising at least one CpG position shall be taken

to mean that all cytosines of said CpG dinucleotides are methylated.

The term "unmethylated" when used in reference to a CpG position shall be taken to mean that the cytosine of said CpG dinucleotide does not carry a methyl group at the 5 position.

The term "methylated" when used in reference to a CpG position shall be taken to mean that the cytosine of said CpG dinucleotide carries a methyl group at the 5 position.

The term "defined methylation status" as used herein shall refer to the methylation status of the cytosine of a CpG dinucleotide wherein said status shall either be methylated or unmethylated at the 5-position.

In the method according to the invention, the DNA to be investigated is hybridized to oligonucleotides of a defined methylation status and then reacted with specific restriction enzymes. The restriction enzymes are able to recognize CpG sequences and also to distinguish hemi-methylated DNA double strands either from unmethylated or from homogenously methylated DNA double strands. These restriction enzymes are denoted in the following as *hemi-methylation sensitive*.

Hybrids of the oligonucleotides and genomic DNA may form two different conformations either hemi-methylated and homogenously methylated. If cytosines which are methylated at the CpG position to be investigated form double strands with oligonucleotides whose corresponding CpG position is unmethylated, in these double strands this particular CpG position is methylated on one strand only this conformation is hemi-methylated. The same applies to hybrids of the corresponding methylated oligonucleotides and unmethylated investigated DNA . In contrast, if both the oligonucleotide as well as the DNA are methylated at the CpG position then a double strand methylated on both strands results homogenous methylation). Vice versa, if the CpG positions both in the DNA to be investigated as well as in the

oligonucleotide are unmethylated, then unmethylated double strands are formed. Thus in the following the terms hemi-methylated, homogenously methylated and unmethylated do not describe the total methylation state of the DNA, but only the state of individual CpG positions within the DNA.

The method of methylation analysis according to the invention is comprised of the following four steps:

a) the DNA to be investigated is hybridized to oligonucleotides of a defined methylation status,

b) the hybrids are reacted with at least one hemi-methylation sensitive restriction enzyme,

c) the occurrence or non-occurrence of a restriction is detected ,

d) the methylation state of the investigated DNA is determined.

In the first step of the method according to the invention, the DNA to be investigated is hybridized to oligonucleotides. This can be conducted both in solution as well as on a solid phase. The DNA to be investigated can originate from different sources. For diagnostic purposes tissue samples among others can be used as the initial material, however body fluids particularly serumcan also be used. It is also conceivable to use DNA from sputum, stool, urine, or cerebrospinal fluid. Preferably, the DNA is isolated from the biological specimens. The DNA is preferably extracted according to standard methods e.g. with the use of the Qiagen UltraSens DNA extraction kit. The isolated DNA is then fragmented e.g. by reaction with conventional (not hemi-methylation sensitive) restriction enzymes. The reaction conditions and the enzymes employed are known to the person skilled in the art e.g. from the protocols supplied by the manufacturers.

The oligonucleotides at the cytosine positions which are recognized by the restriction enzyme that will be used later are unmethylated or methylated at

the 5-position, depending on the restriction enzyme that will be used. For a quantitative analysis, both methylated as well as unmethylated oligonucleotides are utilized (see below). The synthesis of correspondingly unmethylated and methylated oligonucleotides is part of the prior art. In another preferred variant, several oligonucleotides of different sequences are used, so that the simultaneous investigation of several methylation positions is possible.

In a preferred embodiment the oligonucleotides also carry at least one detectable label. A plurality of suitable labels is known to the person skilled in the art. Thus e.g. dyes, fluorescent labels, radionuclides, electrical charge carriers or labels that can be detected in the mass spectrometer may be utilized. Peptide labels are also suitable, which may be detected indirectly by the binding of another labeled antibody. Chemical labels are also suitable, which are detected by subsequent reaction with another labeled marker molecule. Many other labeling options also belong to the prior art. Preferably, the oligonucleotides of different sequences or different methylation statuses carry different labels.

In a particularly preferred variant, the oligonucleotide is labeled with a fluorescent dye on one side of the restriction site and a so-called "quencher" on the other side. If a restriction occurs the dye and the quencher are separated, therefore the dye signal can be detected (Fig. 1). Dyes and quenchers that can be used are known to the person skilled in the art.

The oligonucleotides are preferably bound to a solid phase. The type of solid phase and the solid phase coupling are known in the art. For example, solid phases can include but are not limited to functionalized polymers, metals, glass or semiconductors such as silicon. The oligonucleotides can be bound e.g. via bifunctional linker molecules which are bound to a silanized surface or for example, via thioates or thiol modifications of the oligonucleotide to bromacetyl (= acetyl bromide)-derivatized surfaces or gold. The oligonucleotides of different sequences or different methylation status are

preferably spatially separated from one another so that a separate detection of each oligonucleotide is possible.

In another preferred embodiment, the oligonucleotides are introduced onto a sensitive surface, whose physical or chemical properties can be modified in a measurable way by means of a restriction. For example conductivity, characteristic frequency or surface tension are such measurable properties. In a particularly preferred variant, the surface is comprised of a piezoelectric crystal. The binding of DNA to piezoelectric crystals is known to the person skilled in the art (for review: Skladal: Piezoelectric Quartz Crystal Sensors Applied for Bioanalytical Assays and Characterisation of Affinity Interactions. J. Braz. Chem. Soc., Vol.14, No.4, 491-502, 2003).

The hybridization of the oligonucleotides to the DNA to be investigated is performed under stringent or moderately stringent conditions.

In the second step of the method according to the invention, the hybrids are contacted with hemi-methylation sensitive restriction enzymes under conditions suitable for restriction digest. The selection of the restriction enzymes is made according to the sequence specificity of the enyzmes and according to the diagnostic or scientific objective to be investigated. Enyzmes which preferably cleave unmethylated and hemi-methylated DNA as opposed to homogenously methylated DNA are preferably utilized. If one starts with methylated oligonucleotides in this case then methylated and unmethylated cytosine positions can be different in the DNA to be investigated. The methylated DNA forms homogenously methylated hybrids with the methylated oligonucleotides, and these hybrids are not cleaved by the restriction enzyme. In contrast, in the case of unmethylated DNA, hemi-methylated hybrids are formed, which are recognized by the restriction enzyme and are cleaved. On the other hand, if one starts with unmethylated oligonucleotides then hemi-methylated hybrids are formed with the methylated DNA and unmethylated hybrids with unmethylated DNA. Both hybrids are cleaved by the enzyme so that a differentiation between

methylated and unmethylated DNA is not possible in this case with the use of unmethylated oligonucleotides. The use of unmethylated oligonucleotides in addition to the use of methylated oligonucleotides of the same sequence however, permits a quantification of the method. The ratio of total DNA to methylated DNA can be calculated from the ratio of the two signals. Such a quantification is easily possible e.g., if methylated and unmethylated oligonucleotides are provided with different labels or if they are bound spatially separately on a solid phase (Fig. 2). The person skilled in the art knows how to obtain information on enzymes that can be used in this embodiment. In particular, the REBASE database (currently accessible at http://rebase.neb.com/) provides a great deal of information on hemi-methylation sensitive restriction enzymes. The use of the following enzymes is preferred: AcsII; AdeI; AscI; HinPI; ClaI; EciI; HinP1I; Hpy99I; NruI; RsrII; SalI. The restriction sites of these enzymes are listed in the Appendix. Reaction conditions for the enzyme reaction are known in the art and can be taken e.g. from the protocols supplied by the manufacturers.

In another embodiment of the method according to the invention, enzymes are utilized which preferably cleave unmethylated DNA as opposed to hemi-methylated and homogenously methylated DNA. In this case the use of unmethylated oligonucleotides is necessary (or additionally the application of methylated oligonucleotides for quantification). More detailed information on enzymes that can be utilized is available from the source above.

As long as corresponding enzymes are available, it is in principle also conceivable to work with enzymes which do not cleave unmethylated DNA, but cleave hemi-methylated and homogenously methylated DNA, or which do not cleave unmethylated and hemi-methylated DNA, but cleave homogenously methylated DNA.

It is obvious that biologically active fragments or modifications of the enzymes can also be utilized in the method according to the invention. With the increasing success of enzymatic design, the use of enzymes constructed

especially for the purposes of this invention are also conceivable.

According to the invention, several different restriction enzymes can also be utilized simultaneously or sequentially in combination with different oligonucleotides, in order to investigate the methylation state of several different cytosine positions.

The detection is conducted in the third step of the method according to the invention. This step is carried out by means of techniques known in the art. If labeled oligonucleotides are utilized, then the labels either of the uncleaved oligonucleotides or of the restriction fragments can be detected. When a solid phase is used it is also possible to detect restriction fragments which are found in solution, or fragments which are bound to the solid phase (e.g., with the use of a quencher). By coupling the oligonucleotides to a sensitive surface, the chemical or physical properties that change due to the restriction are measured.

Then in the fourth step, the methylation status of the DNA is concluded from the detected signal and the proportion of methylated DNA is determined.

If disease-specific cytosine positions are investigated, then the method according to the invention is particularly suitable for the diagnosis of cell proliferative disorders (including cancer) or other diseases associated with a change of methylation status. These include, among others, CNS malfunctions; symptoms of aggression or behavioral disturbances; clinical, psychological and social consequences of brain damage; psychotic disturbances and personality disorders; dementia and/or associated syndromes; cardiovascular disease, malfunction and damage; malfunction, damage or disease of the gastrointestinal tract; malfunction, damage or disease of the respiratory system; lesion, inflammation, infection, immunity and/or convalescence; malfunction, damage or disease of the body as a consequence of an abnormality in the development process; malfunction, damage or disease of the skin, the muscles, the connective tissue or the

bones; endocrine and metabolic malfunction, damage or disease; headaches or sexual malfunction. The method according to the invention is also suitable for predicting undesired drug effects and for distinguishing cell types or tissues or for investigating cell differentiation.

The subject of the invention is also the use of hemi-methylation sensitive restriction enzymes for the methylation analysis and for the detection of the above-named diseases associated with a change of methylation status, In a particularly preferred embodiment enzymes selected from the group consisting: AcsI; AdeI; AscI; HinPI; ClaI; EciI; HinP1I; Hpy99I; NruI; RsrII; SalI are utilized for the above-named purposes.

A preferred embodiment of the invention is comprised of a solid surface referred to as a "test strip", on which a plurality of oligonucleotides each of which varies from at least one other oligonucleotide in terms of methylation status and/or sequence are immobilized. This oligonucleotides of the test strip are hybridized to the DNA to be investigated in a temperature-controlled container, preferably a chamber or mini-chamber. Restriction and detection are then performed in one step, e.g., in a cuvette in which the absorption spectra of the dyes used are measured (Fig. 3). In another preferred embodiment, the restriction fragments diffuse to another phase and detectable secondary reactions are therein carried out.

The subject of the present invention is also a kit, comprised of at least one oligonucleotide comprised of one or more CpG positions each of a defined methylation state, at least one methylation sensitive restriction enzyme and suitable buffers.

The subject of the present invention is also a kit, comprised of different immobilized oligonucleotides, at least one hemi-methylation sensitive restriction enzyme and the necessary restriction buffers.

The subject of the present invention is also a kit, comprised of at least one

oligonucleotide comprised of one or more CpG positions each of a known methylation state, at least one methylation sensitive restriction enzyme and suitable restriction buffers.

Example

The following example will illustrate the application of the method according to the invention for characterizing tumors. A tumor develops into two different types (A and B), each of which requires a different treatment. The two types however, are not diagnosable solely on the basis of morphological features; they may be distinguished however, by their methylation status; a CpG base pair which lies within the base sequence GCGC in the middle of a known sequence context is methylated in tumor Type A, whereas it is present unmethylated in Type B. For investigation, DNA from the tumor tissue is extracted using a commercially available kit. A thermal denaturation of the DNA and a subsequent hybridization with an exact 1:1 mixture of the two synthetic oligomers C and D are performed. These latter oligomers possess the same base sequence (complementary to the DNA to be investigated), but are distinguished by their methylation status. The CpG which corresponds to the CpG whose methylation is to be determined is unmethylated in oligomer C but methylated in oligomer D. Oligomers C and D are each is dual-labeled with fluorescent "reporter" and "quencher" dye molecules. Each of said dies is covalently linked to the oligomer such that the "quencher" dye is insufficient proximity to the "reporter" dye to quench fluorescence (and thus detection) of the reporter dye. Restriction of the oligomer between the two labeling sites resulting in the spatial separation of the two dyes leads to the detectable fluorescence of the reporter dye. After the hybridization of the oligomers to the DNA a restriction of the DNA oligonucleotide hybrids that have formed is conducted. For this purpose, a large quantity of the hemi-methylation sensitive restriction enzyme HinPI is added to the DNA. This enzyme cleaves hemi-methylated and unmethylated DNA, but not homogenously methylated DNA in the sequence context GCGC. The increase in fluorescence is

measured over time during the progression of the reaction by means of a temperature-controlled fluorescence photometer. Only the analysis of methylated DNA from the tumor tissue with methylated oligomers, thus in this case Tissue type A hybridized to oligonucleotide type D, results in methylated DNA on both strands of the hybrid DNA structure. This DNA is not cleaved by the restriction enzyme, accordingly no fluorescence is detected. The proportion of methylated DNA in an investigated tumor specimen can be determined by the ratio of the fluorescence of the labels of the two oligonucleotides D and C. This value permits information to be obtained on the type of tumor and thus an optimal treatment.

Brief description of the figures

Fig. 1 shows a preferred embodiment of the invention with the use of a reporter/quencher dual-labeled oligonucleotide . In this case, a methylated oligonucleotide is bound to a solid phase. The oligonucleotide are each labeled with a reporter dye (pentagon) and a quencher (square). The DNA to be investigated is hybridized to the oligonucleotides. It is then reacted with a restriction enzyme. If the DNA to be investigated is unmethylated, then the hybrid is cleaved. Reporter dye and quencher are separated and a signal can be detected.

Fig. 2 shows a preferred embodiment of the invention. In this case, two different types of oligonucleotides are used. Both oligonucleotides have the same base sequence comprising one CpG position. Said CpG position has a different methylation status in each of said oligonucleotides and each is labeled with different dyes (circle and pentagon, respectively). The DNA to be investigated is hybridized to the oligonucleotides. A restriction is then conducted. The degree of methylation of the specimen (M) can be determined from the ratio of the dye signals.

Fig. 3 shows schematically a preferred embodiment of the invention. In this case a tumor specimen is investigated for its methylation status by means of

a test strip. For this purpose, the DNA extracted from the specimen is hybridized to oligonucleotides fixed on a test strip (step 1). A restriction is conducted with a hemi-methylation sensitive restriction enzyme (step 2). Different colored patterns (A, B and C) result, each one dependant on the methylation status of the DNA to be investigated (step 3). It is known from earlier experiments which methylation patterns are associated with specific types of tumor (4). Information concerning appropriate tumor therapies can be derived by comparison with these data.

Appendix: Hemi-methylation sensitive restriction enzymes

The enzymes listed below in Table 1are a selection of suitable enzymes that can be used in the method according to the invention. The restriction sites are taken from the REBASE database. Each of said enzymes is accompanied by figures showing different methylation states of the restriction site (hereinafter referred to as 'conformations'), each of said conformations being presented in one of three columns (left, right and center). The conformations presented on the left column are cleaved at the optimal rate of the restriction reaction. The restriction reaction rate is greatly reduced in the conformations of the center column and does not proceed in the conformations in the right hand column.

Table 1

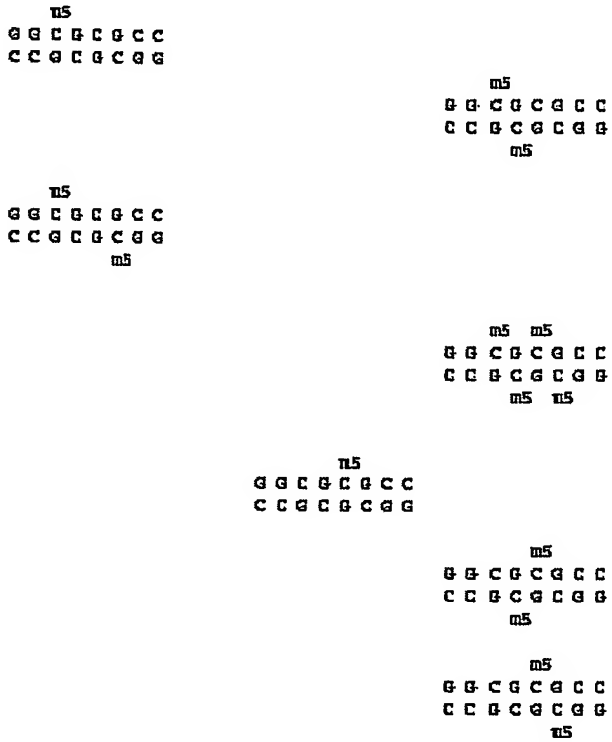
AscI

```
                    C C G C
                    G G C G
                      m5

          m5
C C G C
G G C G

                m5
                C C G C
                G G C G
                  m5

                    m5
                C C G C
                G G C G
```

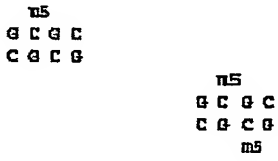
AdeI

```
          m5
C A C N N N G T G
G T G N N N C A C

                    m5
                C A C N N N G T G
                G T G N N N C A C
                          m5
```

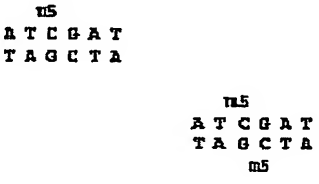
AhdI



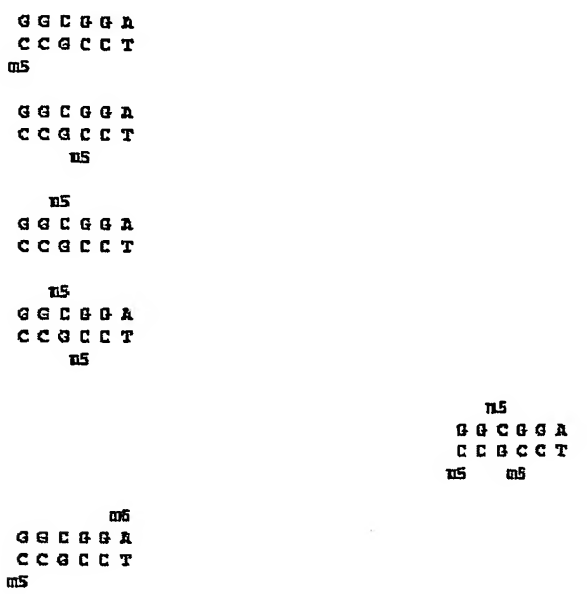
HinPI



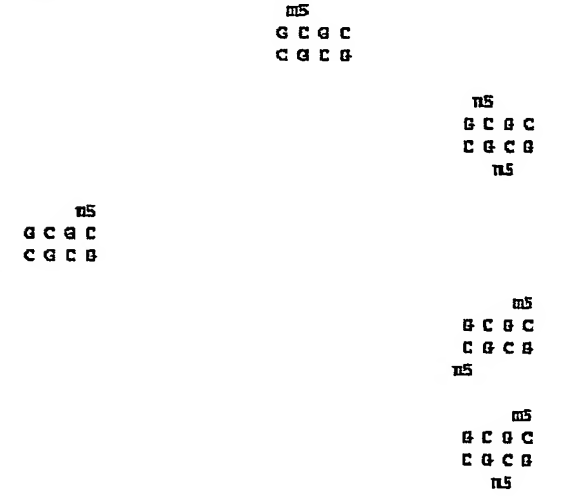
ClaI



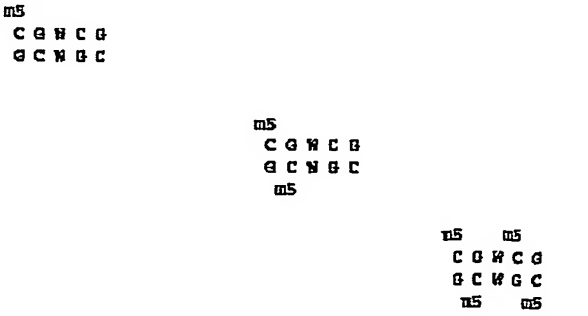
DclI



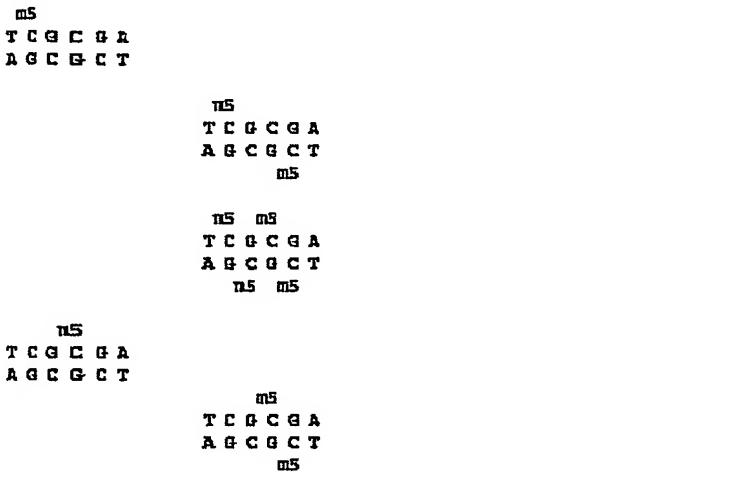
HinPlI



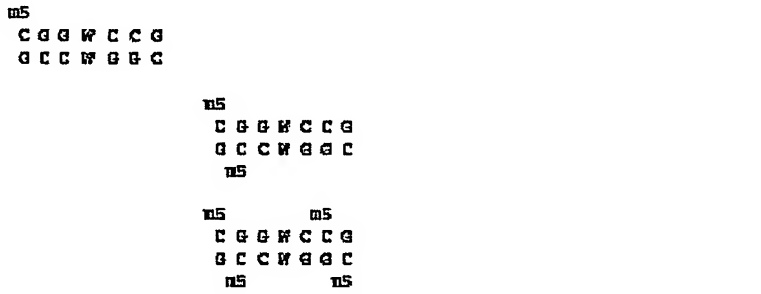
Hpy99I



NruI



BarII



SalI